

Posudek bakalářské práce

předložené na Matematicko-fyzikální fakultě
Univerzity Karlovy v Praze

posudek vedoucího

Autor/ka: Jiří Maršík

Název práce: Rychlý a trénovatelný tokenizér pro přirozené jazyky

Studijní program a obor: Informatika, obor Obecná informatika

Rok odevzdání: 2011

Jméno a tituly vedoucího/oponenta: RNDr. Ondřej Bojar, Ph.D.

Pracoviště: ÚFAL MFF UK

	e x c e l e n t n í	o d p o v í d a j í c í	s l a b š í	n e v y h o v u j í c í
Náročnost zadaného tématu		X		
Míra splnění zadání	X			
Rozsah práce	X	X		
Struktura textové části práce	X	X		
Analýza	X			
Vývojová dokumentace	X	X		
Uživatelská dokumentace	X	X		
Jazyková a typografická úroveň	X	X		
Návrh a design implementace		X		
Kvalita zpracování softwarové části	X	X		
Stabilita aplikace		X		

Nejvýznamnější klady:

- samostatnost při práci, široký záběr, obsáhlá rešerše, výborná práce s externími nástroji a literaturou
- pilotní experimenty a následný přechod od knihoven pro regulární výrazy ke generovaným lexerům; chytrá práce vyžadující pouze 16 módů lexeru
- vynikající celkový výsledek: skutečně rychlý a použitelný nástroj
- konfigurovatelnost
- velmi pečlivé vyhodnocení jak přesnosti, tak rychlosti

Nejzávažnější nedostatky:

Žádné.

Další poznámky:

- Jak probíhá trénování, když vstupní rysy pro klasifikátor zahrnují i seznam posledních rozhodnutí? Jde o výsledky z předchozí iterace trénování?

	v ý b o r n ě	v e l m i d o b ř e	d o b ř e	n e p r o s p ě l / a
Návrh známky	X			

Datum: 21.8.2011

Podpis: